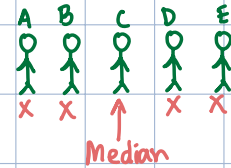


STATISTICS

(25-30 marks) (P1+P2)
(EASIEST PORTION)



- ✓ 1) Mode - Median - Mean - Range
- ✓ 2) Histograms
- 3) Cumulative Frequency
- ✓ 4) Correlation, Scatter Plot.

MODE - MEDIAN - MEAN - RANGE

RAW DATA (NO FREQUENCY TABLE)

Ages of children in a class:

10	8	7	8	10	6	11	10
10	5	7	6	8	9	10	10

Total = n = 16

1] MODE: (MOST OCCURRING VALUE) = 10

2] MEDIAN: MIDDLE VALUE

Step 1: ARRANGE IN INCREASING ORDER

5, 6, 6, 7, 7, 8, 8, 8, 9, 10, 10, 10, 10, 10, 10, 11
X X X X X X X X X X X X X X X

$$\frac{8+9}{2} = \frac{8+9}{2}$$

Median = 8.5

$$\boxed{3} \text{ MEAN : AVERAGE} = \frac{\text{Sum}}{\text{No}}$$

$$= \frac{5+6+6+7+7+8+8+8+9+10+10+10+10+10+10+11}{16}$$

$$= \frac{136}{16} = \boxed{8.5 = \text{Mean}}$$

$$\boxed{4} \text{ RANGE : MAX} - \text{MIN}$$

$$= 11 - 5$$

$$= 6$$

Q. 13, 5, 8, 4, 8, 3, 8 $n=7$

$$\boxed{1} \text{ MODE : } 8$$

$$\boxed{2} \text{ MEDIAN :}$$

STEP 1: ARRANGE 3, 4, 5, 8, 8, 8, 13

x, x, x, ↑, x, x, x

$$\text{Median} = 8$$

$$\boxed{3} \text{ MEAN} = \frac{3+4+5+8+8+8+13}{7} = \frac{49}{7} = 7$$

$$\boxed{4} \text{ RANGE} = \text{Max} - \text{Min} = 13 - 3 = 10$$

MEDIAN:

FIRST METHOD (CUTTING) CAN ONLY USED FOR RAW DATA.
It cannot be used to find median on Table Data.

So, Lets learn a new method for median to be used on table data later on.

(using method 2)

Q: 13, 4, 12, 8, 6, 2, 10, 12 $n=8$

Find median.

Step 1: Arrange: 2, 4, 6, 8, 10, 12, 12, 13

Step 2: Median = $\left(\frac{n+1}{2}\right)^{\text{th}}$ term

$$\left(\frac{8+1}{2}\right)^{\text{th}} \text{ term} = 4.5^{\text{th}} \text{ term}$$

$$\begin{array}{c} \swarrow \quad \searrow \\ \text{4th term} \quad \text{5th term} \\ 8 \quad + \quad 10 \\ \hline 2 \end{array} = 9 = \text{Median}$$

NOTE: CUTTING METHOD IS USED FOR RAW DATA

FORMULA IS USED FOR TABLE DATA.

2 TABLE DATA (SINGLE VALUE TABLE)

x/marks	f	fx = f x x
0	20	20 x 0 = 0
1	10	10 x 1 = 10
2	30	30 x 2 = 60

$0,0,0 \dots$ 20 times
 $1,1,1 \dots$ 10 times
 $2,2,2 \dots$ 30 times

Sigma
 $\sum$ = Add Total
 ↓
 Summation

$n = 60$
 $\sum f = 60$
 $\sum fx = 70$
 Total

Caution: Be careful! Do not write 30 as Mode.

(a) MODE (MODAL VALUE) = 2

(b) MEDIAN: (FORMULA METHOD)

$$\text{Median} = \left(\frac{n+1}{2} \right)^{\text{th}} \text{ term} = \left(\frac{60+1}{2} \right)^{\text{th}} \text{ term}$$

$$= 30.5^{\text{th}} \text{ term}$$

30th term 31st term

$$\frac{1}{2} + \frac{2}{2} = 1.5 = \text{Median}$$

(c) MEAN: For table make a new column: fx

$$\text{Mean} = \frac{\sum fx}{\sum f} = \frac{70}{60} = 1.1667$$

Total (for $\sum fx$)
 Total (for $\sum f$)

Caution: For range we subtract x-values. Do not subtract the frequency values

(d) RANGE: $2 - 0 = 2$

③ TABLE DATA (RANGE TABLE)

Add $\div 2$

$a < x \leq b$
↓
class or interval.

mid value(x)	x/marks	f	$fx = f \times x$
$\frac{5+15}{2} = 10$	$5 < x \leq 15$	20	$20 \times 10 = 200$
$\frac{15+35}{2} = 25$	$15 < x \leq 35$	30	$30 \times 25 = 750$
$\frac{35+40}{2} = 37.5$	$35 < x \leq 40$	10	$10 \times 37.5 = 375$
		$n = 60$ $\Sigma f = 60$	$\Sigma fx = 1325$

(a) MODAL CLASS : (MODE) = $15 < x \leq 35$

(b) MEDIAN INTERVAL.

$$\text{Median} = \left(\frac{n+1}{2} \right)^{\text{th}} \text{term} = \left(\frac{60+1}{2} \right)^{\text{th}} \text{term} = 30.5^{\text{th}} \text{term}$$

$\swarrow$ $30^{\text{th}} \text{ term}$ $\searrow$ $31^{\text{st}} \text{ term}$

Median: $15 < x \leq 35$

Note: For a range table the both values for median WILL ALWAYS BE on the same floor.

(c) MEAN: Make new columns: ① fx
② Mid value of x .

$$\text{Mean} = \frac{\Sigma fx}{\Sigma f} = \frac{1325}{60} = 22.0833 \dots$$

d) RANGE = $40 - 5 = 35$

- 11 The number of items bought by 10 customers at a local store is shown below.

6 7 8 9 10 7 18 10 7 9

(a) State the mode of this distribution. = 7

(b) Find the median number of items bought.

Step 1: Arrange: 5 6 7 7 7 9 9 10 10 18
 ↓
 $\frac{7+9}{2} = 8 = \text{Median}$

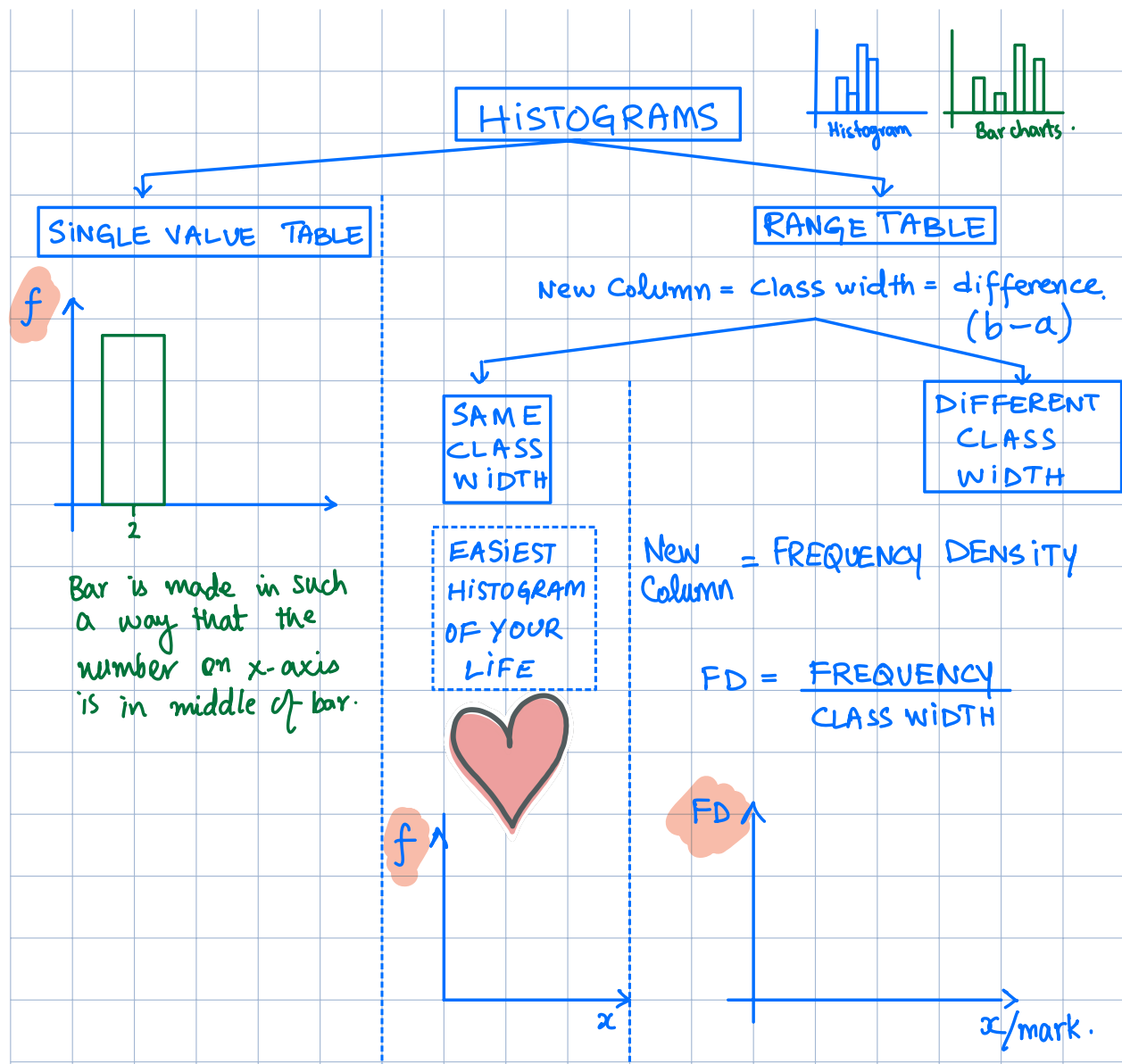
- (b) The table summarises the playing times of each of the 100 tracks on Tom's MP3 player.

Mid(x)	Playing time (t minutes)	Frequency	$fx = f \times x$
3	$2.5 < t \leq 3.5$	5	$5 \times 3 = 15$
4	$3.5 < t \leq 4.5$	30	$30 \times 4 = 120$
5	$4.5 < t \leq 5.5$	50	$50 \times 5 = 250$
6	$5.5 < t \leq 6.5$	15	$15 \times 6 = 90$

Calculate an estimate of the mean playing time of the individual tracks.

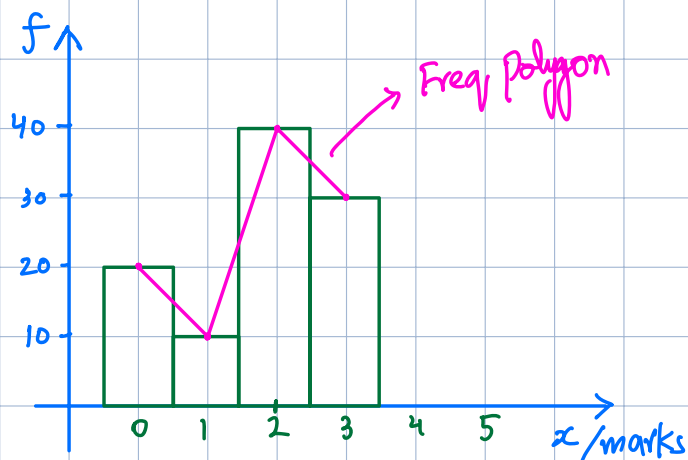
$$\text{Mean} = \frac{\sum fx}{\sum f} = \frac{475}{100} = 4.75$$

Answer (b) minutes [3]



1 SINGLE VALUE TABLE .

x-axis	y-axis
x/marks	f
0	20
1	10
2	40
3	30



NOTE: In single value table if there is a bar at zero on x-axis, move the whole x-axis one box forward.

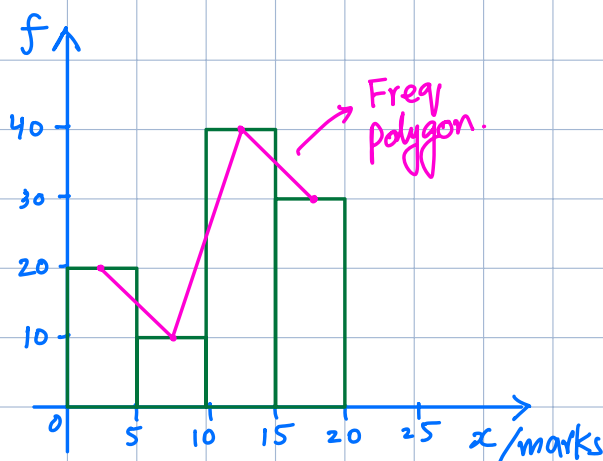
2 RANGE TABLE (SAME CLASS WIDTH) $a < x < b$

STEP 1: Make new column = class width = $b - a$ (Difference)

	x-axis	y-axis
CLASS WIDTH DIFFERENCE: Big-Small	x/marks	f
$5-0=5$	$0 < x \leq 5$	20
$10-5=5$	$5 < x \leq 10$	10
$15-10=5$	$10 < x \leq 15$	40
$20-15=5$	$15 < x \leq 20$	30



If all of these are same, there is no use of this column anymore. You can draw the EASIEST HISTOGRAM OF YOUR LIFE.

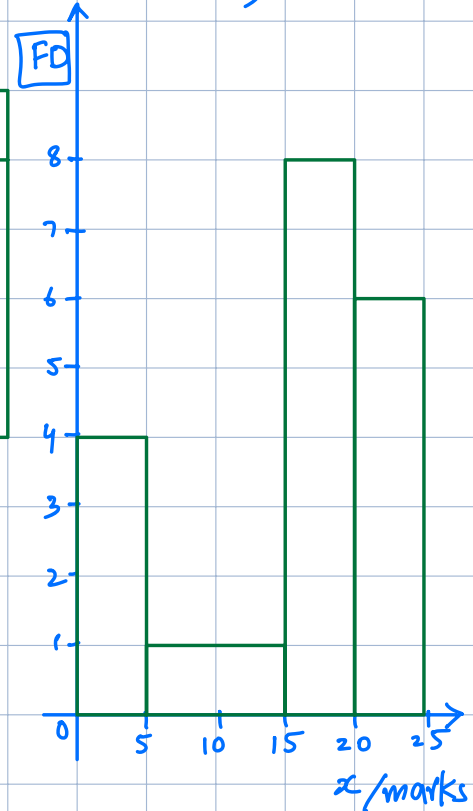


3 RANGE TABLE (DIFFERENT CLASS WIDTH)

	x-axis		y-axis
CLASS WIDTH	x/marks	f	FD = $f \div CW$
5-0 = 5	$0 < x \leq 5$	20	$20 \div 5 = 4$
15-5 = 10	$5 < x \leq 15$	10	$10 \div 10 = 1$
20-15 = 5	$15 < x \leq 20$	40	$40 \div 5 = 8$
25-20 = 5	$20 < x \leq 25$	30	$30 \div 5 = 6$

↓
if all of these are not same, now make a new column for FREQUENCY DENSITY

$$FD = \frac{\text{Frequency}}{\text{CLASS WIDTH}}$$

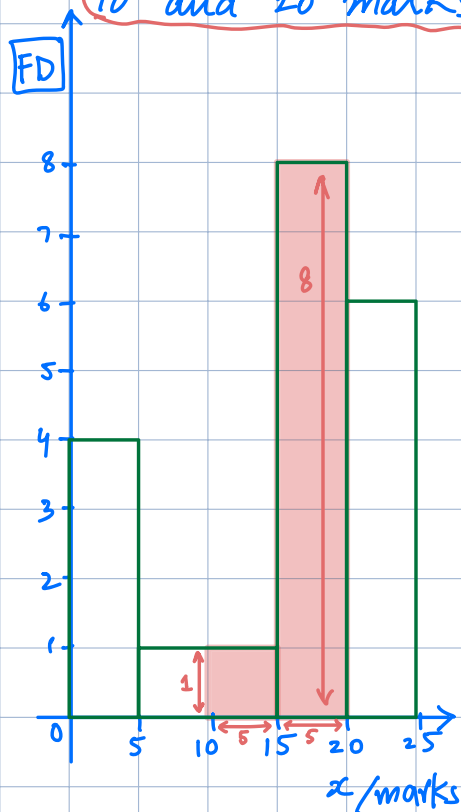


AFTER YOU MAKE A FREQUENCY DENSITY HISTOGRAM.

$$\left[\begin{array}{c} \text{AREA UNDER} \\ \text{GRAPH} \end{array} \right] \text{ OF } \left[\begin{array}{c} \text{FREQUENCY} \\ \text{DENSITY} \\ \text{HISTOGRAM} \end{array} \right] = \left[\begin{array}{c} \text{NUMBER} \\ \text{OF THINGS} \end{array} \right]$$

Area under graph

Q. Find number of students who scored between 10 and 20 marks?



$$(1 \times 5) + (5 \times 8) \\ 5 + 40 \\ = 45 \text{ students.}$$

CUMULATIVE FREQUENCY (12 MARKS)

(TOTAL)

One day, garage A records the amount of petrol bought by the first 120 customers.
The results are summarised in the table below.

FREQUENCY TABLE:

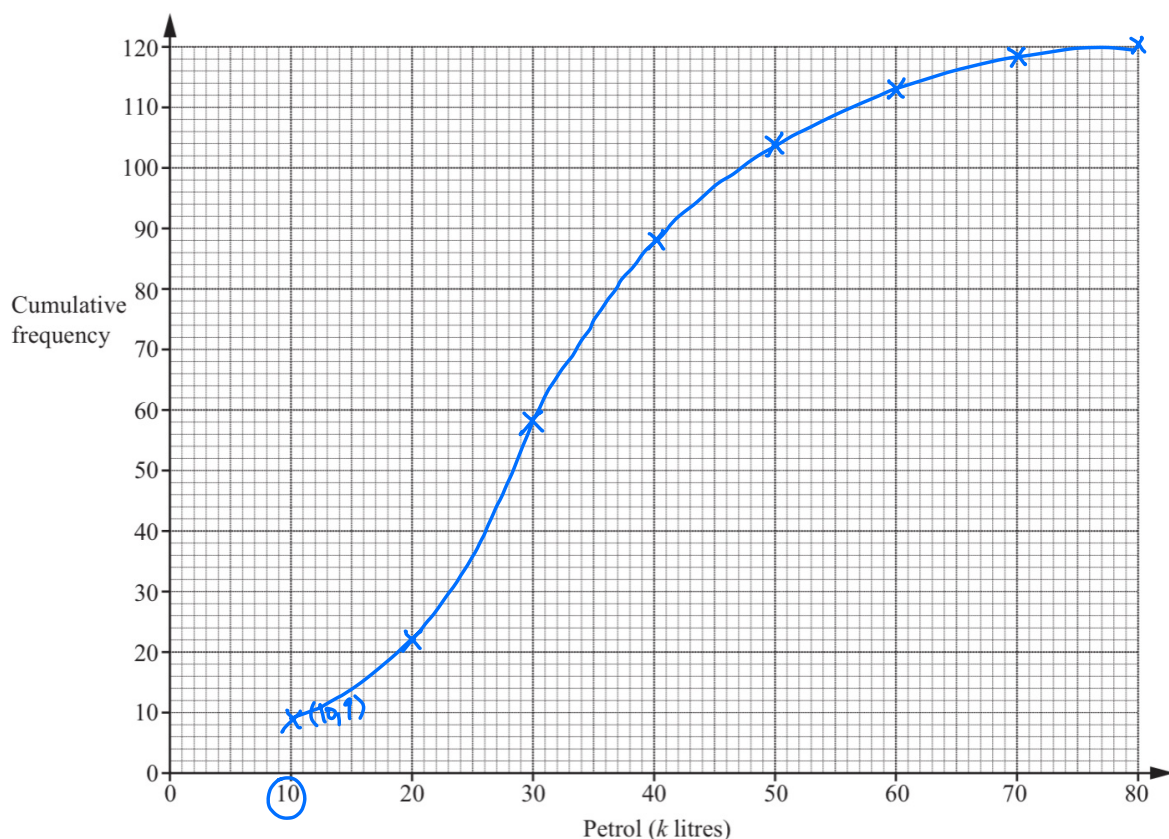
Petrol (k litres) (x)	$0 < k \leq 10$	$10 < k \leq 20$	$20 < k \leq 30$	$30 < k \leq 40$	$40 < k \leq 50$	$50 < k \leq 60$	$60 < k \leq 70$	$70 < k \leq 80$
Number of customers (f)	9	13	36	30	16	9	5	2

(a) Complete the cumulative frequency table below.

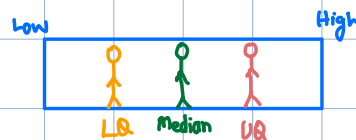
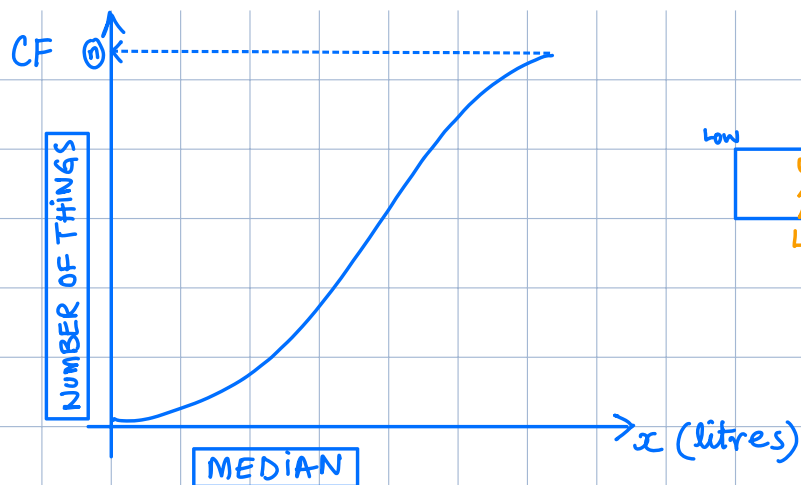
Petrol (k litres)	$k \leq 10$	$k \leq 20$	$k \leq 30$	$k \leq 40$	$k \leq 50$	$k \leq 60$	$k \leq 70$	$k \leq 80$
(TOTAL) Cumulative frequency	9	22	58	88	104	113	118	120

(10, 9) (20, 22) (30, 58) (40, 88) (50, 104) (60, 113) (70, 118) (80, 120)

(b) On the grid below, draw a cumulative frequency curve to represent this data.



[3]



MEDIAN

UPPER QUARTILE

LOWER QUARTILE

INTER QUARTILE RANGE

PERCENTILE

X-AXIS

Caution Formula
for median is
 $(\frac{n+1}{2})^{\text{th}}$ term every
where except
cumulative freq.
where median = $\frac{n}{2}$

1) $\text{MEDIAN} = \frac{n}{2} = \frac{120}{2} = \cancel{60}$ check x-axis for MEDIAN $\rightarrow$ Median = 30.5

2) $\text{UPPER QUARTILE} = \frac{3n}{4} = \frac{3(120)}{4} = \cancel{90}$ check x-axis for UQ $\rightarrow$ UQ = 41

3) $\text{LOWER QUARTILE} = \frac{n}{4} = \frac{120}{4} = \cancel{30}$ check x-axis for LQ $\rightarrow$ LQ = 23

4) $\text{INTER QUARTILE RANGE} = \text{upper Quartile} - \text{Lower Quartile}$
 $= 41 - 23$
 $= 18$

Note: Make sure to subtract x-axis values of UQ and LQ.

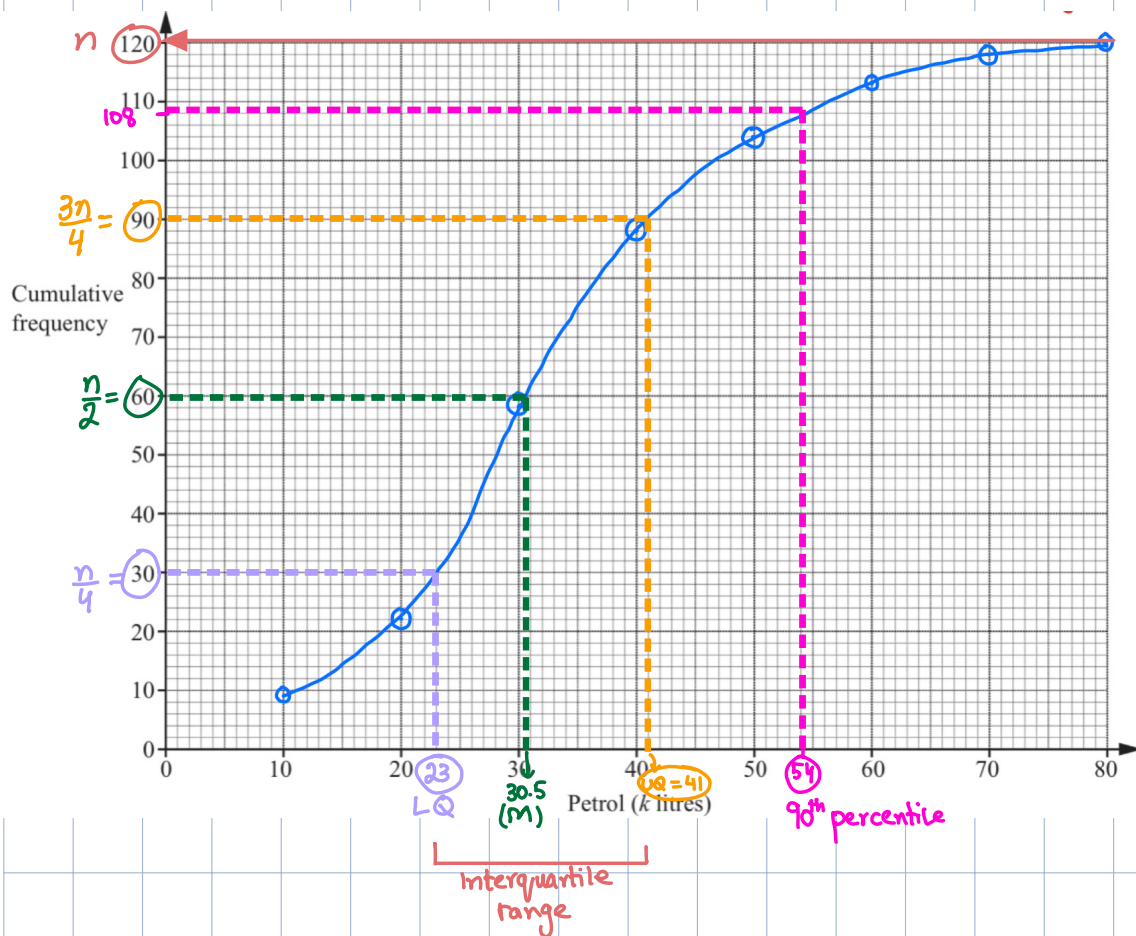
5) **PERCENTILE** (For k^{th} Percentile) = $n \times \frac{k}{100}$ = check x-axis.

For example if you have to find 90th percentile:

Question will tell you which percentile to find.

$$120 \times \frac{90}{100} =$$

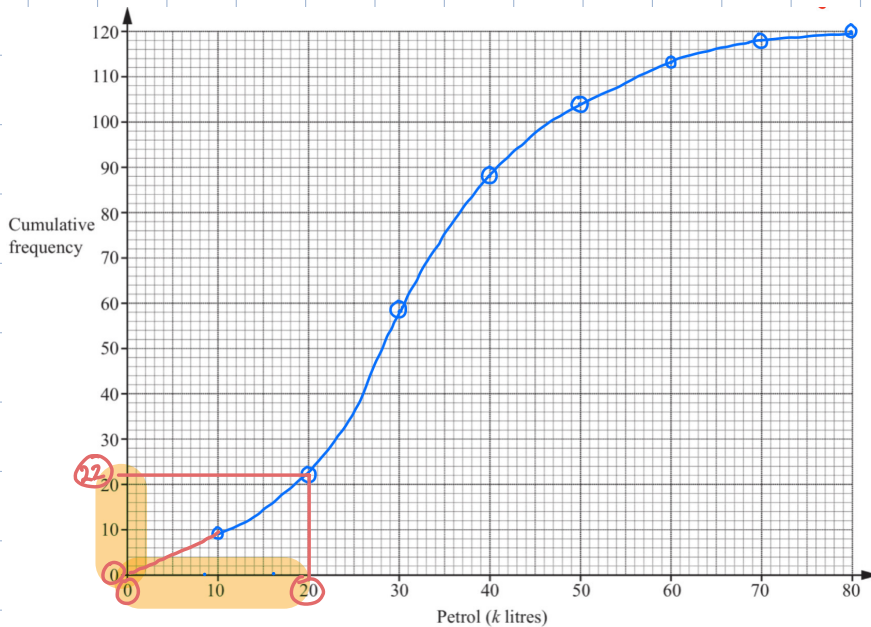
~~108~~
check x-axis



(ALWAYS MAKE QUESTION SENTENCE "BETWEEN — and —")

Y-AXIS : NUMBER OF THINGS.

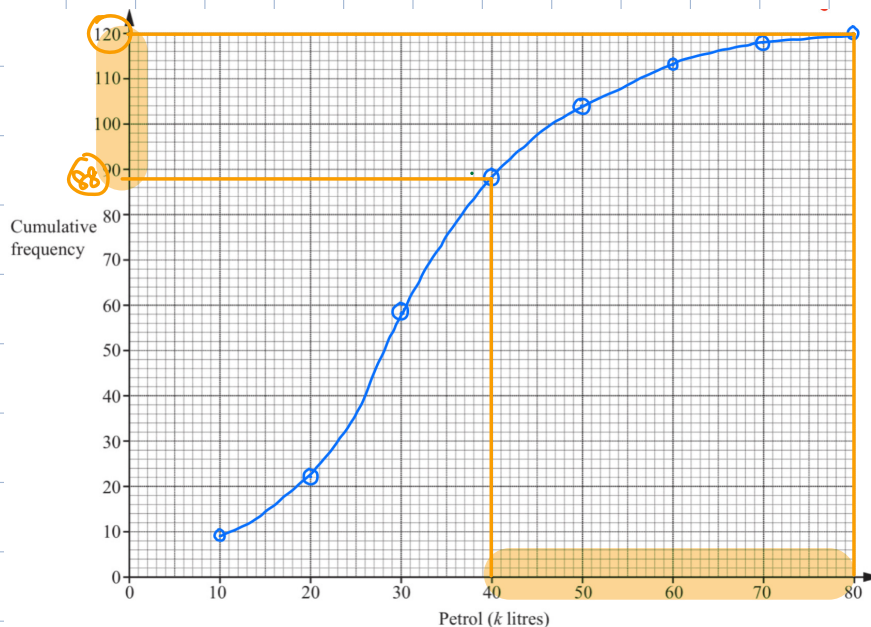
- ii) Find number of cars who bought less than 20 litres.



between 0 and 20
↓ ↓
0 22

$$22 - 0 \\ = 22 \text{ cars.}$$

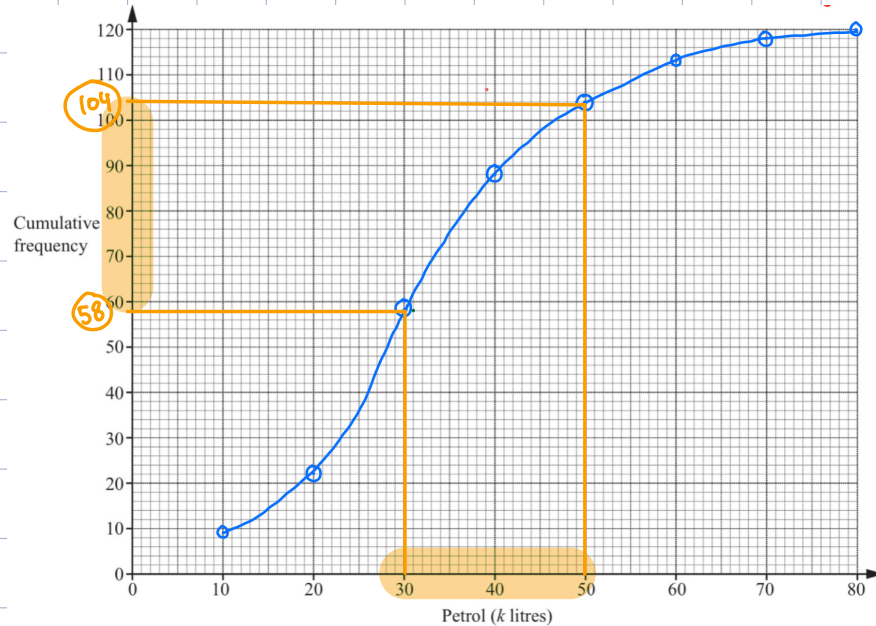
- iii) Find number of cars who bought more than 40 litres.



between 40 and 80
↓ ↓
88 120

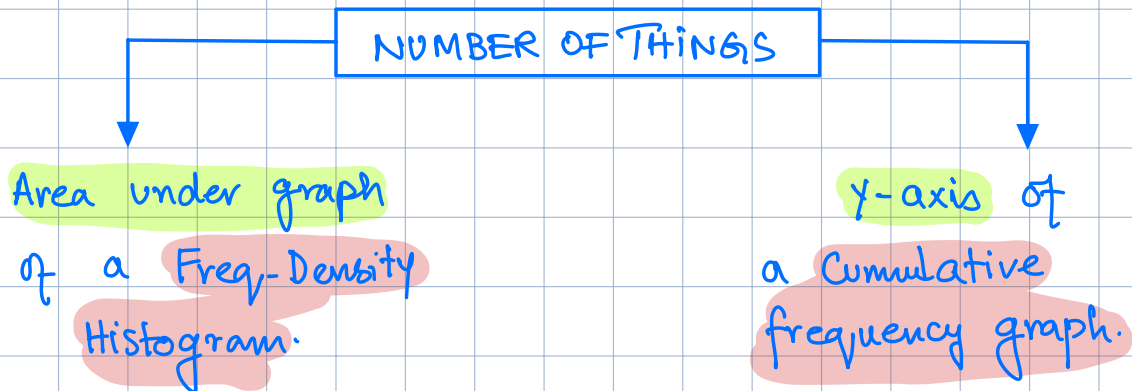
$$120 - 88 \\ = 32 \text{ cars}$$

(iii) Find number of cars who bought between 30 and 50 litres.



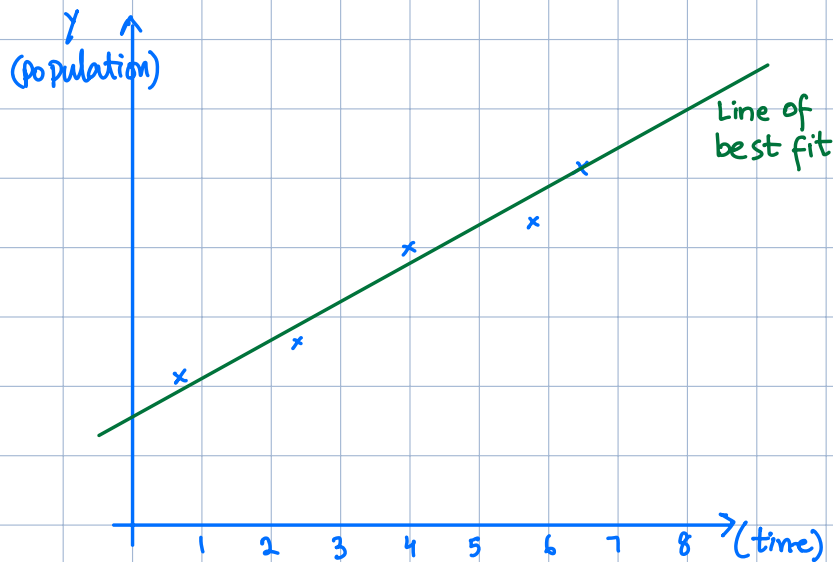
between 30 and 50
 $\downarrow$ 58 $\downarrow$ 104

$$104 - 58 = 46 \text{ cars}$$



SCATTER PLOTS

(LINE OF BEST FIT) (CORRELATION).



Concept

Job of line of best fit is to show trend of plots

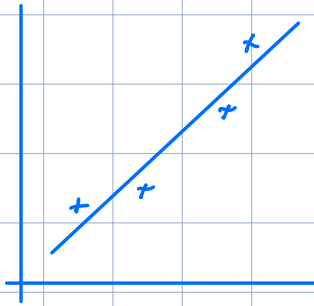
LINE OF BEST FIT:
(TREND LINE)

DO'S

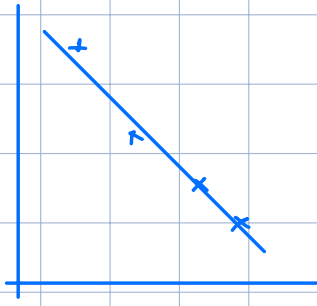
- 1) Average line through the points.
- 2) Almost equal balance of points on each side of line.
- 3) Line of best fit must represent TREND of points

DONT'S

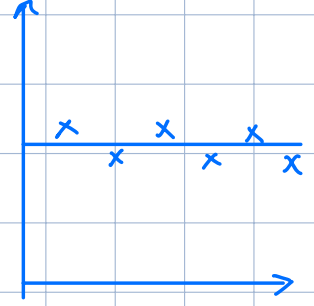
- 1) NEVER FORCE THE LINE THROUGH ORIGIN (0,0)
- 2) NEVER FORCE THE LINE THROUGH FIRST AND LAST POINT.
- 3) NEVER FORCE LINE TO TOUCH MAXIMUM POINTS.



POSITIVE
CORRELATION



NEGATIVE
CORRELATION



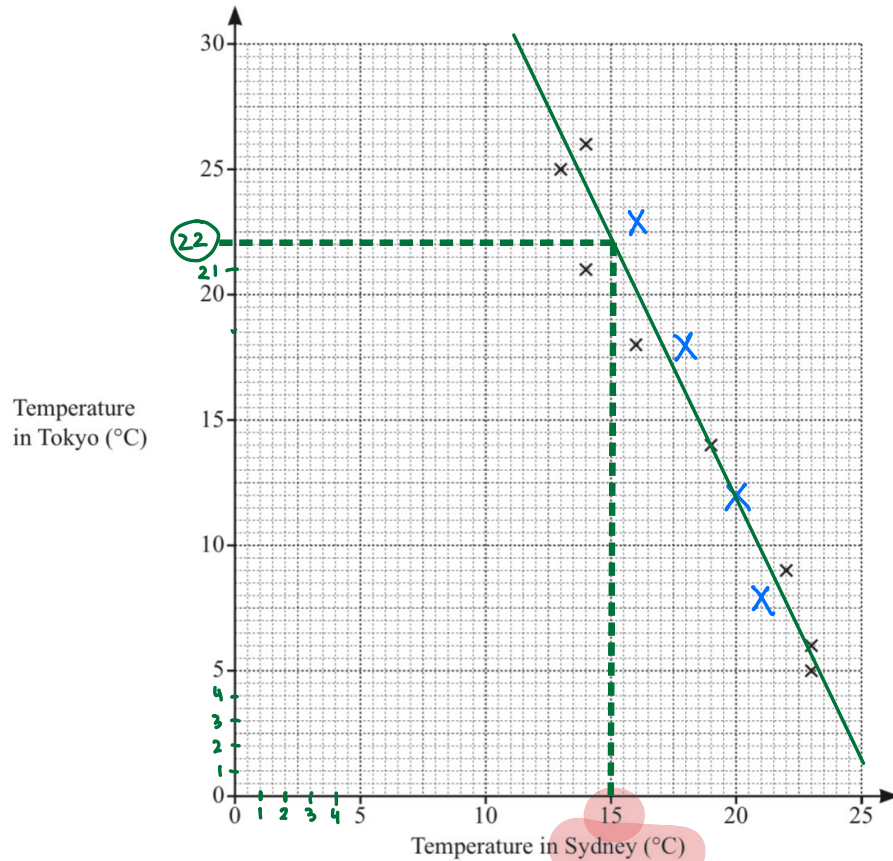
NO (Zero)
CORRELATION

- 2 The table shows the average monthly temperatures ($^{\circ}\text{C}$) in Tokyo and in Sydney one year.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Temperature in Sydney ($^{\circ}\text{C}$)	23	23	22	19	16	14	13	14	16	18	20	21
Temperature in Tokyo ($^{\circ}\text{C}$)	5	6	9	14	18	21	25	26	23	18	12	8

- (a) Complete the scatter diagram.
The first eight points have been plotted for you.

[2]



- (b) What type of correlation is shown by the scatter diagram?

Negative [1]

- (c) Draw a line of best fit.

[1]

- (d) The following year, the average temperature in Sydney during May was 15°C .

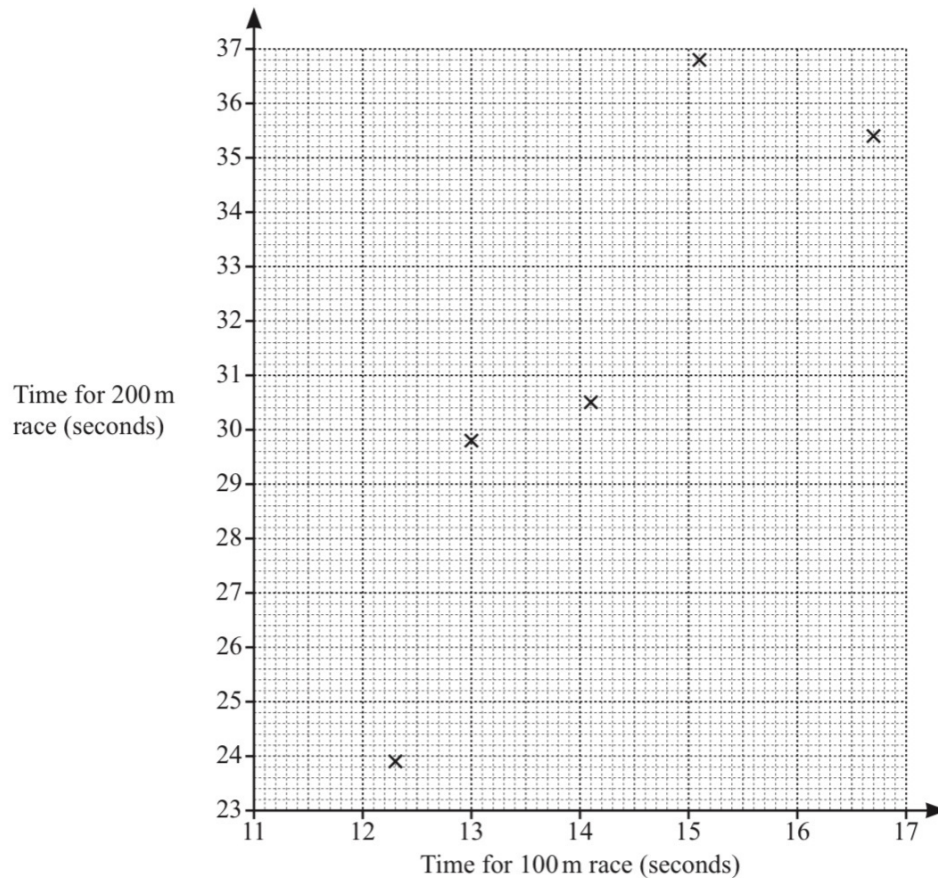
By using your line of best fit, estimate the average temperature in Tokyo that May.

y-axis 22 $^{\circ}\text{C}$ [1]

- 2 Ten boys ran in a 100 m race and a 200 m race.
The table below shows their times in seconds.

Time for 100 m race	12.3	14.1	15.1	16.7	13.0	14.7	13.7	12.9	15.2	16.1
Time for 200 m race	23.9	30.5	36.8	35.4	29.8	32.5	28.4	26.1	33.5	36.0

- (a) Complete the scatter diagram.
The first five points have been plotted for you.



[2]

- (b) What type of correlation is shown in the scatter diagram?

..... [1]

- (c) Draw a line of best fit.

[1]

- (d) Another boy recorded a time of 27.5 s in the 200 m race.

Use your graph to estimate the time it would take him to run 100 m.

..... s [1]

STATS TALLY

one
two
three
Four
Five

Nine
Thirteen

Q: 5, 4, 5, 5, 4, 2, 2, 1, 3 Ages.

Age	Tally	Frequency.
1	/	1
2	//	2
3	/	1
4	//	2
5	///	3

5
Ⓜ

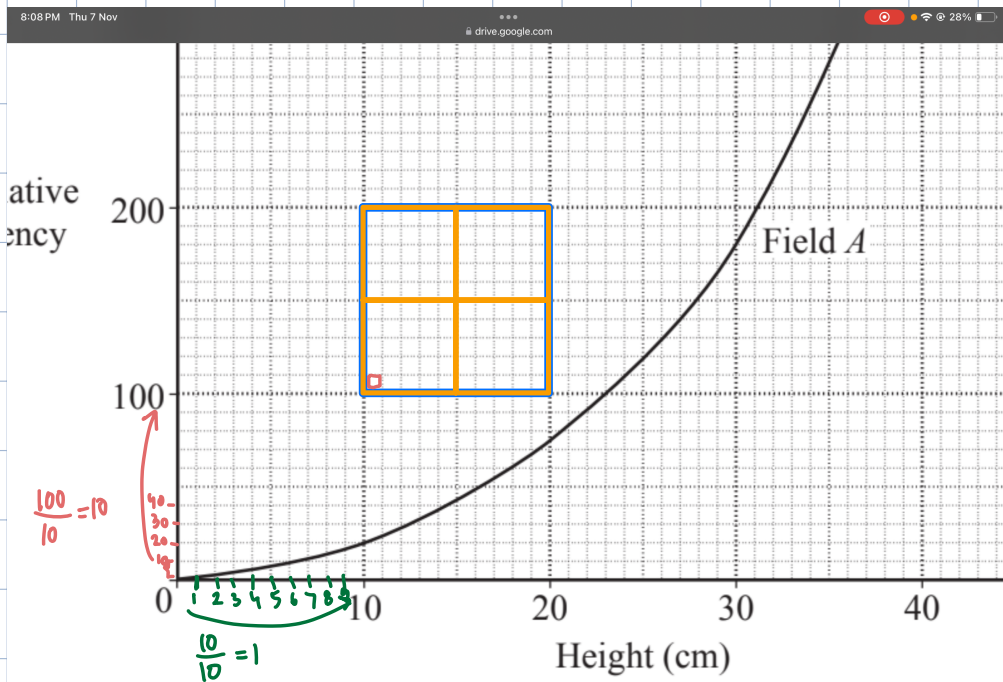
8 = Ⓜ ///

14 = Ⓜ Ⓜ ////

Pictogram

 = 10000 cars

Months	Sale	Value
Jan	  	30000
Feb		5000
March	  	25000



GUIDE FOR NEW COLUMNS

